

FERRAMENTAS PARA ANÁLISE DE EFICÁCIA DE SISTEMAS INTERATIVOS DE RECUPERAÇÃO DA INFORMAÇÃO.

Nilson Augusto Magalhães Moraes Neto¹; Rodrigo Tripodi Calumby²

1. Bolsista PROBIC/UEFS, Graduando em Engenharia da Computação, UEFS, e-mail: nilson.uefs@gmail.com

2. Orientador, Departamento de Ciências Exatas, UEFS, e-mail: rtcalumby@ecomp.uefs.br

PALAVRAS-CHAVE: Recuperação da Informação; Avaliação de Eficácia; Busca Interativa.

INTRODUÇÃO

Dado o avanço e popularização das tecnologias de captura e armazenamento de informação, um conjunto grande de conteúdos digitais tem sido criado. A necessidade de explorar de forma eficiente estas bases de conteúdo é apresentada em várias áreas, como: ciências biológicas e da natureza, medicina, redes sociais, recomendação à usuário, etc. Desta forma, a construção de métodos eficazes para recuperação e indexação destas bases é imprescindível.

Uma ferramenta, chamada de AnalyzIR, tem sido desenvolvida na UEFS (Silva, 2014) com o intuito de auxiliar o processo de análise de eficácia desses métodos. Atualmente a ferramenta realiza a avaliação de eficácia para um ou vários sistemas com base em resultados médios ou por consultas específicas, como pode ser visto no trabalho de Silva et al (2014), além do comparativo entre eles, utilizando-se de um conjunto extenso de medidas disponíveis.

Atualmente, sistemas no quais o usuário interage, como por exemplo, realizando feedbacks aos resultados retornados estão cada vez mais presentes e são chamados de iterativos. Existem algumas dificuldades em comparar resultados iterativos. Há um desafio em conciliar os métodos baseados no comportamento e cognição do usuário. As buscas reais são muito complexas e envolvem objetos que não são capturados pelas medidas utilizadas atualmente.

Neste trabalho, a ferramenta em questão foi aprimorada e expandida, tendo sido desenvolvidas as funcionalidades de avaliação desses sistemas.

METODOLOGIA

Não existem métodos experimentais padronizados para avaliação de eficácia destes sistemas, sendo assim, as avaliações são feitas através de protocolos. Por conta dessas adversidades, o *trec_eval* (principal ferramenta amplamente utilizada para calcular eficácia de sistema de recuperação da informação) (TREC, 2014) não é capaz de naturalmente analisar e comparar sistemas iterativos, apesar do programa ter sido inicialmente projetado para executar esta tarefa, durante seu desenvolvimento observou-se adversidades na comparação entre esses sistemas, principalmente na dificuldade entre analisar os resultados devido à falta de padronização na execução dos experimentos. Os arquivos utilizados para o cálculo possuem um campo inoperante que inicialmente foi projetado para ser identificar a iteração, a permanência do campo (sempre em com 0) foi devido a construção de diversos scripts que já tinham sido configurados, ou seja, foi muito melhor manter o campo do que reconfigurar todos os scripts já implementados. A ferramenta é capaz de analisar iterativamente por meio de um pré-processamento dos dados onde os arquivos de entrada (Figura 1 e 2) são formatados com o mesmo padrão utilizado pelo *trec_eval* (utilizando o campo previamente inativo) e re-organizados de acordo a iteração.

Os protocolos utilizados para avaliação dos sistemas iterativos são: *rank-shift*, no qual os itens relevantes indicados pelo usuário em uma dada iteração são realocados no topo do ranking nas iterações subsequentes; *collection reranking*, onde todas as amostras são

reranqueadas à cada interação; *residual collection*, onde somente os itens não-vistos são apresentados ao usuário nas iterações seguintes; e *full-freezing* no qual as relevantes são mantidas no mesmo lugar e ao final da sessão de busca é montado um ranking com a concatenação das amostras vistas pelo usuário à cada interação na mesma ordem que elas foram vistas.

RESULTADOS

Um projeto iterativo é constituído por dois tipos de arquivos de entrada: os resultados dos sistemas a serem analisados (runs) (Figura 1) e um gabarito utilizado pelo trec_eval para calcular a eficácia do sistema (qrels) (Figura 2). Além dos arquivos, o usuário deve indicar qual protocolo foi utilizado para a obtenção dos resultados. Será apresentado o passo a passo para a criação de um projeto iterativo e geração de gráficos (Figura 4 à 6). A depender da configuração de gráfico que o usuário seleciona, a ferramenta monta um gráfico diferente que melhor se adeque à análise, tendo um total de seis tipos de gráficos diferentes (Tabela 1). A figura 7 foi gerada a partir dos parâmetros selecionados na Figura 5, já a Figura 8 foi construída com os seguintes parâmetros: odas as 3 runs; a média de todas as consultas; todas as 10 iterações; medida recall com profundidade 10.

Sistemas	Consultas	Iterações	Medidas
1	vários	1	1
vários	vários	1	1
1	1	vários	1
vários	1	vários	1
vários	1	1	1
1	1	1	vários
vários	1	1	vários
1	vários	vários	1
1	vários	1	vários
1	1	vários	vários

Tabela 1. Combinações de parâmetros selecionáveis. Cada cor representa um tipo de gráfico.

1	1	A24	1	89	my_run
1	1	A30	2	88	my_run
1	1	A10	3	87	my_run
1	1	A27	4	86	my_run
1	1	A14	5	85	my_run
1	1	A7	6	84	my_run

Figura 1. Formato de arquivo de run. Atributos em cada coluna (nessa ordem): identificador da consulta, iteração, identificador do arquivo (nome), ranque, escore, nome da run.

1	0	A2	0
1	0	A1	1
2	0	A30	1
2	0	A29	0
2	0	A28	1

Figura 2. Formato do arquivo qrels. Atributos em cada coluna (nessa ordem): identificador da consulta ,
 iteração (sempre em zero), identificador do arquivo, relevância.

Figura 3: Tela para a criação de um novo projeto iterativo. É solicitado ao usuário um nome para um projeto, os diretórios dos resultados e do gabarito além de qual protocolo foi utilizado.

Figura 4. Primeira tela para criação de gráfico. Nela o usuário tem a possibilidade de escolher
 quais parâmetros serão utilizados para gerar os gráficos.

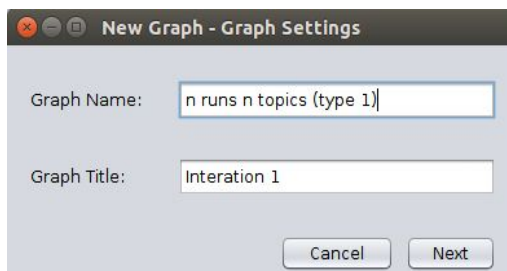


Figura 5. Seleciona nome e título do gráfico.

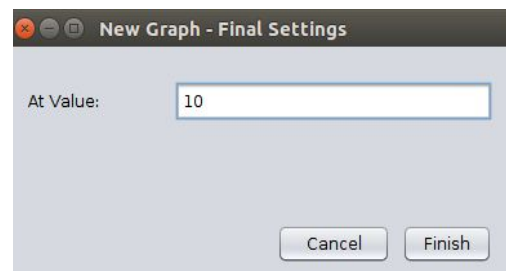


Figura 6. Seleciona a profundidade do rank que será avaliada.

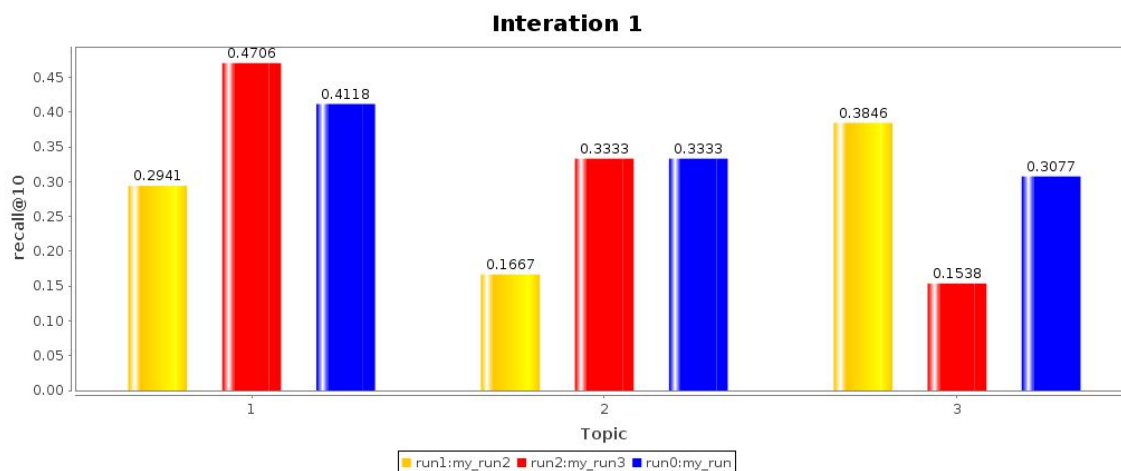


Figura 7. Gráfico obtido de acordo a parametrização selecionada pelo usuário.

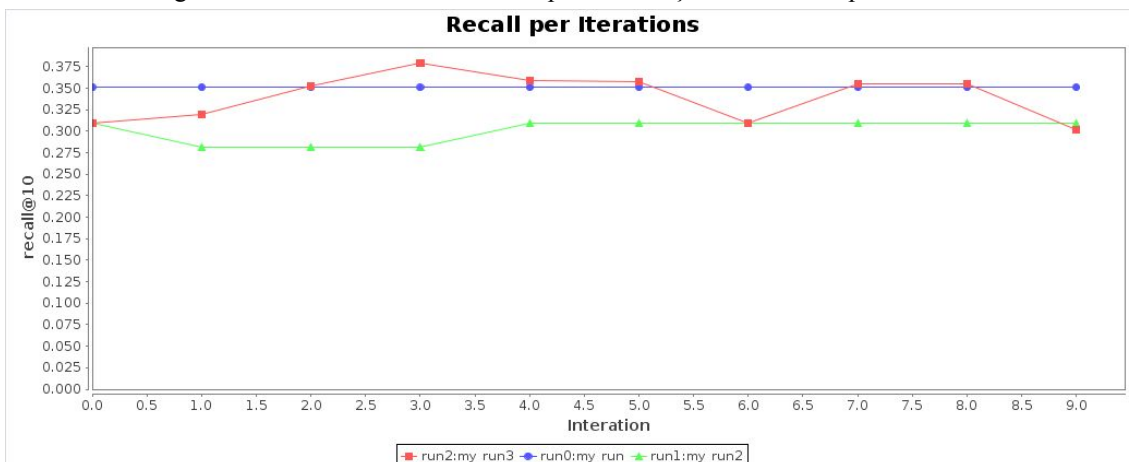


Figura 8. Gráfico obtido do mesmo projeto criado na Figura 3.

CONSIDERAÇÕES FINAIS

O resultado obtido é algo importante e inovador no campo de pesquisa da recuperação de informação, pois possibilita que o trec_eval seja utilizado para calcular resultados por iteração, função que naturalmente a ferramenta não executa, incorporando os diferentes protocolos de avaliação. Com isso, a AnalyzIR ficou mais robusta, tornando mais fácil o usuário avaliar sistema iterativos e possibilitando uma melhor interpretação dos resultados adquiridos com determinado sistema.

REFERÊNCIAS

SILVA, L.A; MAIA, A.L.M., CALUMBY, R.T. 2014. Uma plataforma extensível para análise de resultados de eficácia em Recuperação de Informação. In: XVIII Seminário de Iniciação Científica da UEFS, Feira de Santana.

TREC. 2014 [online]. Notes on Trec Eval. HomePage:http://ir.iit.edu/~dagr/cs529/files/project_files/trec_eval_desc.html